Infinite Sets

Course notes from NAIL063 by doc. Kynčl

Winter semester 2025/2026

Jakub Smolík

smolikj@matfyz.cz

Contents

1 Review of Set Theory Basics

We will be dealing with the Zermelo–Fraenkel (ZF) set theory; after adding the axiom of choice, we obtain the much stronger ZFC theory. Let us first recall some important axioms and definitions.

As for notation, I always use the symbol $\subset$ for a proper subset (or subclass) and $\subseteq$ for a general subset (or subclass). I use $\subsetneq$ only when it is important that the two sets (or classes) are not equal. Concatenated expressions such as $a \in b \in c$ mean $a \in b \wedge b \in c$. I differentiate between the symbol for equality of two objects "=" and the symbol for definition of an object ":=". Similarly, I use "$\Longleftrightarrow$" for equivalence in logical formulas and "$\equiv$" for defining concepts by an equivalent statement. For example: the relation R is symmetrical $\equiv R = R^{-1}$. I use the following notation for defining functions.

- $f : A \to B$ is a function with domain A and codomain B.
- $f : a \mapsto b$ denotes that f maps the set $a \in A$ to the set $b \in B$.

I use the terms function, map, and mapping interchangeably.

1.1 Sets vs. Classes

Definition 1.1 (Class)**.** If $\varphi(x)$ is a formula, then the expression $\{x \,|\, \varphi(x)\}$ is called a *class term*. It defines the "collection" of all sets x satisfying $\varphi(x)$. We call this collection the *class* determined by $\varphi(x)$.

Every set is a class but not all classes are sets (take the class of all sets). A class that is not a set is called a *proper class*. The major difference between sets and classes is that classes cannot be members of other classes or sets, while sets can. We can substitute class terms into logical formulas in place of free variables, but unlike sets, we cannot quantify them using $\forall$ and $\exists$. It isn't hard to show that for every formula with class terms (but without quantified class variables), there is an equivalent formula in the base language without class terms.

We will usually denote sets using small letters $a, b, c, x, y, \ldots$ and classes using capital letters A, B, C, etc. The exception to this are well-ordered sets which will often be denoted as W. Finally, the class of all sets, also called the *universal class*, is denoted by V.

1.2 Axiom Schema of Replacement

Axiom 1.2. When we take any (even a class) map F and a preimage set a, then the class of images $b = F[a]$ is also a set. Formally, if $\psi(x,y)$ is a formula, without free variables y_1, y_2 and b, then the formula

$$(\forall x)(\forall y_1,y_2)\big((\psi(x,y_1) \wedge \psi(x,y_2)) \Rightarrow y_1 = y_2\big) \Rightarrow \\ (\forall a)(\exists b) : (\forall y)\big(y \in b \Leftrightarrow (\exists x)(x \in a \wedge \psi(x,y))\big)$$

is an axiom. The formula $\psi(x,y_1)$, resp. $\psi(x,y_2)$ are created from $\psi(x,y)$ by substituting y_1, resp. y_2 for y.

The first part of this axiom says that $\psi(x,y)$ should behave like a map $y = F(x)$. In the second part, a denotes the set of preimages and b the set of corresponding images.

1.3 Axiom of Choice

The *axiom of choice* (AoC) is one of the most important principles in modern mathematics, with profound implications in areas such as analysis or linear algebra. It states that for any collection of nonempty sets, it is possible to choose exactly one element from each set, even if the collection is infinite. When added to Zermelo–Fraenkel set theory, it yields the much more powerful ZFC. Many theorems that seem intuitively true, such as every vector space having a basis, depend on this axiom.

However, the axiom of choice is also controversial, since it leads to counterintuitive results, such as the well-ordering principle, which claims that every set can be well-ordered, or the Banach–Tarski paradox, which provides a way to decompose a solid ball into finitely many pieces and reassembled into two identical copies of the original.

Definition 1.3 (Choice function)**.** A *choice function* (or a *selector*) on the set x is any function $f : x \to \bigcup x$ such that

$$(\forall t \in x)(t \neq \varnothing \Rightarrow f(t) \in t).$$

We can WLOG assume that the choice function is defined on $x \setminus \{\varnothing\}$ and all $t \in \mathrm{Dom}(f)$ satisfy $f(t) \in t$.

Axiom 1.4 (Axiom of Choice)**.** Every set has a choice function.

An equivalent formulation which sounds even more "obviously true" is that the the cartesian product of a nonempty indexed family of nonempty sets is nonempty. Other famous conditions equivalent [1] to the axiom of choice are the well-ordering principle and Zorn's lemma.

Showing the equivalence of the nonempty cartesian product statement isn't hard. The other two conditions are much more difficult and require the use of transfinite induction. We prove their equivalence in Sections 2.4 and 2.5.

1.4 Natural Numbers and the Axiom of Infinity

We use Von Neumann ordinals, meaning that natural numbers are defined as

$$0 \coloneqq \varnothing,\ 1 \coloneqq \{0\},\ 2 \coloneqq \{0, 1\}, \ldots\ n + 1 \coloneqq \{0, 1, \ldots n\} = n \cup \{n\}.$$

Definition 1.5. The *successor function* is a mapping $S : V \to V$ defined as $v \mapsto v \cup \{v\}$. For convenience, we write $v + 1 \coloneqq S(v) = v \cup \{v\}$.

Definition 1.6. We say that a set w is *inductive* $\equiv 0 \in w \ \wedge\ (\forall n \in w)(n + 1 \in w)$.

Axiom 1.7 (Axiom of Infinity)**.** There exists an inductive set.

Definition 1.8. We define the *set of all natural numbers* as the smallest inductive set. Or equivalently, as $\bigcap\{w \mid w \text{ is inductive}\}$. We denote it ω.

[1] "The Axiom of Choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?" — Jerry Bona

1.5 Well-Ordering and Initial Segments

Let us recall a very important definition, the notion of *well-ordered sets.*

Definition 1.9 (Ordering)**.** A binary relation R on the class X is a

- *trichotomy* $\equiv (\forall x,y \in X)(x\,R\,y \vee y\,R\,x \vee x = y)$,
- *strict order* $\equiv$ it is anti-reflexive, strongly anti-symmetric and transitive on X; (note that strong anti-symmetry follows from the other two),
- *(partial) order* $\equiv$ it is reflexive, weakly anti-symmetric and transitive on X,
- *total (or linear) order* $\equiv$ it is a trichotomous partial order on X.

If R is an ordering, then instead of $x\,R\,y$ we write $x \leq_R y$ and we call $(X, \leq_R)$ an *ordered class.* Similarly, if R is a strict ordering, then we write $x <_R y$ and we call $(X, <_R)$ a *strictly ordered class.*

Note that we can easily create a strict ordering $<_R$ from $\leq_R$ and vice versa. For this reason, we will not define properties for both strict and non-strict orderings separately because one implicitly defines the other.

We call an element *minimal* if there is no smaller one, and we call it a *minimum* if it is smaller than all others. Similarly, we define a *maximal* and *maximum* element.

Observation 1.10. *Every minimum is minimal. Furthermore, if R is a total order, then there is at most one minimal element, and if it exists, then it is also the minimum. There is always at most one minimum.*

Definition 1.11 (Well-ordering)**.** An ordered class $(A, \leq_R)$ is *well-ordered* $\equiv$ every non-empty subset of A has a minimum. We can similarly define a strict well-ordering.

Observation 1.12. *If $(X, \leq_R)$ is a well-ordered class and $Y \subseteq X$, then $(Y, \leq_R)$ is also a well-ordered class.*

Note that every well-ordered class is also totally ordered because we can take any two elements, and one of them has to be the minimum and is therefore smaller.

Definition 1.13 (Lower closure)**.** Let $(A, <_R)$ be a (strictly) ordered class. A subclass $X \subseteq A$ is a *lower closure* of $A \equiv (\forall x \in X)(\forall a \in A)(a <_R x \Rightarrow a \in X)$. If additionally $X \neq A$, then we call it a *proper lower closure.*

Lemma 1.14. *Let $(W, <_R)$ be a (strictly) well-ordered set and suppose that the set X is a proper lower closure of W. Then there exists a unique $x \in W$ such that $X = \{y \in W \mid y <_R x\}$. We denote this set as $(\leftarrow, x)$.*

Proof. We define x as the minimum of $W \setminus X$. Then every $y <_R x$ belongs to X, so $(\leftarrow, x) \subseteq X$. We also want the opposite inclusion. For contradiction, suppose there is a $y \in X$ s.t. $y \notin (\leftarrow, x)$. If $y \not< x$, then necessarily $x \leq_R y$. But this means that $x \in X$ because $y \in X$ and X is a lower closure. But this is a contradiction because $x \notin X$. $\square$

Definition 1.15 (Initial segment)**.** If $(W, <_R)$ is a (strictly) well-ordered set, then we call its proper lower closures *initial segments* instead. We denote the unique initial segment of W determined by $x \in W$ as

$$(\leftarrow, x) \coloneqq \{y \in W \mid y <_R x\}.$$

It contains all the elements of W from the minimum of W until x, but not x itself.

Observation 1.16. *Note that $x <_R y \iff (\leftarrow, x) \subset (\leftarrow, y)$.*

2 Ordinal Numbers

Informally, *ordinal numbers* are the generalization of natural numbers. We will first do a quick recap of ordinal number basics and then prove a theorem that deeply links ordinals and well-ordered sets.

2.1 Ordinals as a Generalization of Naturals

Definition 2.1. The class X is *transitive* $\equiv y \in x \in X \implies y \in X$. Or equivalently, if for all $x \in X$ also $x \subseteq X$.

Theorem 2.2. *Every natural number and the set of all natural numbers ω are*

1. *transitive and*
2. *(strictly) well-ordered by the membership relation $\in$.*

Remark. From now on, we will denote the (strictly) well-ordered set $(\omega, \in)$ as $(\omega, <)$ instead and write $n < m$ instead of $n \in m$ when talking about natural numbers.

Definition 2.3 (Ordinal numbers)**.** The set α is an *ordinal number* $\equiv$

1. α is transitive and
2. α is (strictly) well-ordered by the membership relation $\in$.

The class of all ordinal numbers is denoted as On.

Observation 2.4. *Every $n \in \omega$ and ω itself are ordinal numbers.*

Theorem 2.5. *The class* On *itself is transitive and (strictly) well-ordered by $\in$. This implies that it is not a set, because otherwise* On $\in$ On*. Furthermore, any proper class X that is transitive and well-ordered by $\in$ is identical to* On*.*

As for notation, we will use symbols $\alpha, \beta, \gamma, \ldots$ to denote ordinals and the symbol $<$ to compare them. That is we write $\beta < \alpha$ instead of $\beta \in \alpha$.

Observation 2.6. *If $\beta < \alpha$, then $\beta \subset \alpha$ and β is an initial segment of α. Additionally, $\alpha = (\leftarrow, \alpha)$.*

Definition 2.7. If $\alpha \in \mathrm{On}$, then we call all $\beta < \alpha$ the *predecessors* of α. The *successor* of α is the ordinal $\alpha + 1 := \alpha \cup \{\alpha\}$. We say that α is the *direct predecessor* of $\alpha + 1$.

Remark. It is easy to show that $\alpha + 1$ is the smallest ordinal larger than α.

Definition 2.8. The ordinal α is an

- *isolated* (also *successor*) ordinal $\equiv \alpha = 0$ or α has a direct predecessor.
- *limit* ordinal $\equiv$ it is not isolated.

Example. Every $n \in \omega$ or $\omega + 1$ is isolated, but ω is a limit ordinal. In fact, ω is the smallest infinite ordinal.

2.2 Ordinals as Types of Well-Ordered Sets

The definition of ordinals we just saw is one by Von Neumann from 1923. However, Cantor originally defined ordinals in 1895 as types of well-ordered sets. We will prove a theorem linking these two concepts together.

Lemma 2.9. *Every proper lower closure of* $(\mathrm{On}, <)$ *is an ordinal number (and therefore a set).*

Proof. Let X be a proper lower closure of On. Then

1. X is transitive. Suppose $\alpha \in \beta \in X$, that is $\alpha < \beta \in X$. Because X is a lower closure we have $\alpha \in X$.
2. X is well-ordered by $\in$ because On is well-ordered by $\in$ and $X \subseteq \mathrm{On}$.

We also need to argue that X is a set. If it were a proper class, then by Theorem 2.5 it would be the entire On, but $X \subsetneq \mathrm{On}$. □

Definition 2.10 (Isomorphism)**.** Let $(A, \leq_R)$ and $(B, \leq_S)$ be ordered classes. A bijection $F : A \to B$ is an *order-isomorphism* of $(A, \leq_R)$ and $(B, \leq_S) \equiv$

$$(\forall x,y \in A)\big(x \leq_R y \iff F(x) \leq_S F(y)\big).$$

Because we will not be dealing with other types of isomorphisms, we will usually simply say “isomorphism” instead of “order-isomorphism”.

Theorem 2.11 (About comparing well-orderings)**.** *If* $(W_1, \leq_1)$ *and* $(W_2, \leq_2)$ *are well-ordered sets, then exactly one of the following holds:*

1. *either* W_1 *and* W_2 *are isomorphic, or*
2. W_1 *is isomorphic to an initial segment of* W_2*, or*
3. W_2 *is isomorphic to an initial segment of* W_1*.*

In each case, the isomorphism is unique.

Corollary 2.12. *No two distinct ordinals can be isomorphic.*

Proof. Suppose $\alpha < \beta$, that is $\alpha \in \beta$ and $\alpha \subset \beta$. Clearly α is an initial segment of β. Which means that we are in case 2 of the previous theorem. $\square$

Theorem 2.13 (About the type of well-ordering). *Every well-ordered set is isomorphic to a unique ordinal number. This ordinal is called the* type *of the ordering.*

The following proof is taken from [2].

Proof. Let $(W, <_R)$ be a well-ordered set. We want to show that there is a unique ordinal $(\alpha, <)$ isomorphic to it. Define X as the set of all $x \in W$ for which $(\leftarrow, x)$ is isomorphic to an ordinal. As no two distinct ordinals are isomorphic, this ordinal is uniquely determined, and we denote it α_x; we denote the isomorphism as $i_x : (\leftarrow, x) \to \alpha_x$.

Suppose that there exists a set S such that $S = \{\alpha_x \mid x \in X\} \subseteq \mathrm{On}$. Because we assume that S is a set then $S \subsetneq \mathrm{On}$. We claim that S is a proper lower closure of $(\mathrm{On}, <)$, and therefore by Lemma 2.9 it is an ordinal, let's call it α. Indeed, suppose $\beta < \alpha_x \in S$, we want $\beta \in S$. Note that β is an initial segment of α_x. This implies that $i_x^{-1}[\beta]$ is an initial segment of W. Because W is well-ordered, $i_x^{-1}[\beta]$ is equal to $(\leftarrow, b)$ for some $b \in W$, (using Lemma 1.14). So $\beta = \alpha_b \in S$ by the definition of S. More precisely, $i_x \restriction (\leftarrow, b)$ is an isomorphism between $(\leftarrow, b)$ and β. We will argue why we can make the assumption that S is a set later.

A similar argument shows that X is a lower closure of W. To show this, suppose $x \in X$ and take $y \in a$ such that $y <_R x$. We want $y \in X$. We have $y <_R x$, therefore $(\leftarrow, y)$ is an initial segment of $(\leftarrow, x)$. Because isomorphisms conserve all ordering properties, $i_x \restriction (\leftarrow, y)$ is an isomorphism between $(\leftarrow, y)$ and an initial segment of α_x. By Lemma 2.9, this is an ordinal; by our previous notation, α_y. Therefore $y \in X$.

We conclude that either $X = W$ or $X = (\leftarrow, c) \subset W$ for some $c \in W$, (using Lemma 1.14). We now define a function $f : X \to S = \alpha$ by $f : x \mapsto \alpha_x$. From the definition of S and the fact that $x < y$ implies $(\leftarrow, x) \subset (\leftarrow, y)$ and therefore $\alpha_x < \alpha_y$ it is obvious that f is an isomorphism of $(X, <_R)$ and $(\alpha, <)$. If

- $X = (\leftarrow, c)$, then by the definition of the set X, $c \in X$ because $(\leftarrow, c)$ is isomorphic to an ordinal $\alpha_c = \alpha$. But this is a contradiction because $c \notin (\leftarrow, c) = X$.
- Therefore $X = W$ and α is the sought-after ordinal isomorphic to $(W, <_R)$.

The uniqueness of α follows from the simple observation that if W was isomorphic to two distinct α_1 and α_2, then by transitivity of isomorphism, α_1 would be isomorphic to α_2, which is impossible by Corollary 2.12.

This would complete the proof if we were justified to make the assumption that the class S is a set and therefore an ordinal. In fact, we have to use the axiom of replacement to guarantee it. If we assume this axiom then S is a set because it is the image of the set X by the map f. $\square$

Exercise. Is there a well-ordered proper class not isomorphic to $(\mathrm{On}, <)$?

2.3 Transfinite Induction and Recursion

In mathematics, we often use induction on the natural numbers to prove statements, and we can use recursion such as $f(0) = 1$ and $f(n) = n \cdot f(n-1)$ to define functions. We will now show how to generalize this to all ordinals.

Theorem 2.14 (The Transfinite Induction Principle)**.** *Let $A \subseteq \mathrm{On}$ be a class such that for all ordinals $\alpha \in \mathrm{On}$ we have $\alpha \subseteq A \Rightarrow \alpha \in A$, or in other words*

$$(\forall \beta < \alpha)(\beta \in A) \implies (\alpha \in A). \tag{2.1}$$

Then $A = \mathrm{On}$.

Equivalently, if $\varphi(x)$ is a property, and for all ordinals α:

$$\textit{If } \varphi(\beta) \textit{ holds for all } \beta < \alpha, \textit{ then } \varphi(\alpha).$$

Then $\varphi(\alpha)$ holds for all ordinals α.

Proof. Suppose that some ordinal $\gamma \notin A$ and let $S = \{\alpha \leq \gamma \mid \alpha \notin A\}$. Because ordinals are well-ordered, the set S has a minimum element α. Since every $\beta < \alpha$ is in A, it follows by (2.1) that $\alpha \in A$, which is a contradiction.

The equivalence can be easily seen by taking the class $A = \{x \mid \varphi(x)\}$ or the property $\varphi(x) \equiv x \in A$. □

We can also formulate the principle separately for isolated and limit ordinals, which allows us to use the transfinite induction principle in a form closer to the usual formulation of the induction principle for the naturals.

Theorem 2.15 (The Transfinite Induction Principle II.)**.** *Let $A \subseteq \mathrm{On}$ be a class satisfying*

(a) $0 \in A$,

(b) $\alpha \in A \Rightarrow \alpha + 1 \in A$, *. . . this is just induction on ω*

(c) if α is a limit ordinal and $(\forall \beta < \alpha)(\beta \in A)$, then $\alpha \in A$.

Then $A = \mathrm{On}$. Note that we can again easily reformulate this in terms of a property $\varphi(x)$.

Proof. We need to show that these three assumptions imply (2.1). So let α be an ordinal such that $\beta \in A$ for all $\beta < \alpha$. If $\alpha = 0$, then $\alpha \in A$ by (a). If $\alpha \neq 0$ is isolated, i.e., if there is a $\beta < \alpha$ such that $\alpha = \beta + 1$, we know that $\beta \in A$, so $\alpha \in A$ by (b). If α is a limit ordinal, we have $\alpha \in A$ by (c). □

We can use transfinite induction to prove properties of certain infinite structures. On the other hand, so-called transfinite recursion allows us to construct various infinitely complex structures and define functions in a recurrent fashion.

Theorem 2.16 (About construction by transfinite recursion)**.** *If $G : V \to V$ is a class map, then there is a unique class map $F : \mathrm{On} \to V$ satisfying*

$$F(\alpha) = G(F \restriction \alpha). \tag{2.2}$$

So we define the image of the next ordinal using its predecessors and their images.

Remark. This should seem a bit suspicious because it looks like we are saying that for every class G there exists a class F that something holds. But we cannot quantify classes. Well, we can replace the quantification of G by a theorem *schema*, one for every G. And we aren't really quantifying F because the following proof explicitly constructs it.

Remark. The theorem can be equivalently formulated using different recurrences, for example as

- $F(\alpha) = G\big(F[\alpha]\big) = G\big(\{F(\beta) \mid \beta < \alpha\}\big)$,
- $G : \mathrm{On} \times V \to V$ and $F(\alpha) = G(\alpha, F \restriction \alpha)$,
- $F(\alpha)$ is $G_1(F(\beta))$ if $\alpha = \beta + 1$ is isolated, and $G_2(F[\alpha])$ if α is limit.

Additionally, these transfinite recursion statements are equivalent to the axiom of replacement.

Proof. We define A as the class of "set approximations" of F. That is set mappings f, the domain of which is some ordinal number β, and that for all $\alpha < \beta$ we have $f(\alpha) = G(f \restriction \alpha)$. Now we define F as $F \coloneqq \bigcup A$. Clearly $F \subseteq \mathrm{On} \times V$. We will show that $F : \mathrm{On} \to V$ is the unique mapping satisfying (2.2).

First we show that the approximations of F agree. Let f,$f' \in A$ and $\alpha \in \mathrm{Dom}(f) \cap \mathrm{Dom}(f')$. We claim that $f(\alpha) = f'(\alpha)$. Note that $\mathrm{Dom}(f) \cap \mathrm{Dom}(f')$ is an ordinal δ. For contradiction, suppose that $\alpha \in \delta$ is the smallest ordinal for which $f(\alpha) \neq f'(\alpha)$. Then $f \restriction \alpha = f' \restriction \alpha$ so $f(\alpha) = G(f \restriction \alpha) = G(f' \restriction \alpha) = f'(\alpha)$, a contradiction.

Second we verify that F satisfies (2.2), i.e. for all $\alpha \in \mathrm{Dom}(F)$ we have $F(\alpha) = G(F \restriction \alpha)$. So let $\alpha \in \mathrm{Dom}(F)$. It is there due to some $f \in A$ satisfying $\alpha \in \mathrm{Dom}(f)$ and $f(\alpha) = G(f \restriction \alpha)$. Also, $F(\alpha) = f(\alpha)$ and $F \restriction \alpha = f \restriction \alpha$. Therefore, by combining these equalities $F(\alpha) = G(F \restriction \alpha)$.

Next we show that $\mathrm{Dom}(F) = \mathrm{On}$. First we prove that $\mathrm{Dom}(F)$ is a lower closure of On. Suppose $\alpha \in \mathrm{Dom}(F)$, then it is there thanks to some $f \in A$ with domain $\delta > \alpha$. If $\beta < \alpha$, then also $\beta \in \delta$ and thus $\beta \in \mathrm{Dom}(F)$.

According to Lemma 2.9 either $\mathrm{Dom}(F) = \mathrm{On}$, which we want, or $\mathrm{Dom}(F) = \gamma \in \mathrm{On}$. Suppose for contradiction that $\mathrm{Dom}(F) = \gamma$. Then F is a set because $\mathrm{Dom}(F)$ is a set, $\mathrm{Rng}(F)$ is a set using the axiom of replacement, and $F \subseteq \mathrm{Dom}(f) \times \mathrm{Rng}(f)$. This implies that $F \in A$ because its domain is an ordinal and we have verified that it satisfies the recursive definition property.

Now that $F \in A$, we define a slightly "longer" function $F_1 \coloneqq F \cup \{(\gamma, G(F))\}$; note $F = F_1 \restriction \gamma$. Notice that $F_1 \in A$ because $\mathrm{Dom}(F_1) = \gamma + 1$ is an ordinal and we defined it so that it satisfies the recursive definition property. Because $F = \bigcup A$, this implies $F_1 \subseteq F$, but then $\gamma \in \mathrm{Dom}(F_1) \subseteq \mathrm{Dom}(F) = \gamma$, which is a contradiction. We conclude that $\mathrm{Dom}(F) = \mathrm{On}$.

Finally, we prove the uniqueness of F. For contradiction suppose that there is another mapping $F' \neq F$ satisfying this theorem. Because $(\mathrm{On}, <)$ is well-ordered, we can take the smallest ordinal α where $F(\alpha) \neq F'(\alpha)$. Therefore $F \restriction \alpha = F' \restriction \alpha$ and so $F(\alpha) = G(F \restriction \alpha) = G(F' \restriction \alpha) = F'(\alpha)$, which is a contradiction. □

Exercise. Prove by induction on ω that every infinite well-ordered set A, such that each initial segment $(\leftarrow, a)$ is finite, is isomorphic to $(\omega, <)$.

Hint. Since each $(\leftarrow, a)$ is finite, there is a unique $n_a \in \omega$ with the same cardinality. The isomorphism we are looking for is $f : A \to \omega$ defined by $f : a \mapsto n_a$.

Exercise. Prove by transfinite induction that every well-ordered proper class, such that each initial segment $(\leftarrow, a)$ is a set, is isomorphic to $(\mathrm{On}, <)$.

We will use transfinite induction to prove the equivalence of the well-ordering theorem and Zorn's lemma to the axiom of choice. But transfinite recursion can also be used to prove some wildly sounding geometrical claims, such as

- $\mathbb{R}^3$ is a union of pair-wise disjoint unit circles, or that
- there is a set in $\mathbb{R}^2$ that intersects every line in exactly two points.

2.4 The Well-Ordering Principle

The *well-ordering principle* states that every set can be well-ordered. It is also sometimes referred to as the well-ordering theorem or Zermelo's theorem.

Principle 2.17 (Well-Ordering Principle)**.** Every set can be well-ordered.

Theorem 2.18. *The well-ordering principle is equivalent to the axiom of choice.*

Proof. WO $\Rightarrow$ AoC. Let $A \neq \varnothing$ be a set, WLOG $\varnothing \notin A$. We want to construct a selector $f : A \to \bigcup A$ such that for all $a \in A$ we have $f(a) \in a$. The well-ordering principle guarantees a well-ordering $\leq$ on $\bigcup A$ and because every a is a nonempty subset of $\bigcup A$, it has a least element with respect to $\leq$. We chose this minimum as $f(a)$.

AoC $\Rightarrow$ WO. Let $A \neq \varnothing$ be a set. We will use transfinite recursion to label the elements of A by ordinal numbers and then use the well-order of the ordinals to define a well-order on A. Let $g : \mathcal{P}(A) \to A$ be a selector on $\mathcal{P}(A)$, assigning each nonempty $B \subseteq A$ an element $b \in B$. We will want to use transfinite recursion based on g, so we should extend it to be a class map $G : V \to V$, for example by defining it to be equal to $\varnothing$ when g is not defined.

We can now use transfinite recursion to define the function $F : \mathrm{On} \to A \cup \varnothing$ as $F(0) = G(A)$ and $F(\alpha) = G(A \setminus F[\alpha])$. This function assigns each ordinal a unique element from A until they "run out" (when $F[\alpha] = A$), and then it assigns $\varnothing$ to all larger ordinals.

Define W as the class of all ordinals α for which $F[\alpha] \subsetneq A$. Denote the restriction of F to W as $F_W : W \to A$. Plan: show that W is an ordinal, from that F_W is a bijection of W and A and we can denote the unique ordinal mapped to $a \in A$ as α_a. Once we have that, we can define a well-ordering R of A as

$$a <_R b \equiv \alpha_a < \alpha_b.$$

This is a well-ordering because $(A, <_R)$ is order-isomorphic to $(W, <)$, which is well-ordered (because it is an ordinal).

Firstly, we claim that W is a set. Indeed, because F_W is injective, it has an inverse F_W^{-1}, which maps the set $\mathrm{Rng}(F_W) \subseteq A$ onto W which is therefore, using the axiom of replacement, a set. Now we claim that W is a lower closure of On, and so it is an ordinal (by Lemma 2.9). Suppose $\alpha \in W$, i.e. $F[\alpha] \subsetneq A$, and let $\beta < \alpha$. Then $\beta \subseteq \alpha$ and $F[\beta] \subseteq F[\alpha]$, so $\beta \in W$.

To complete the proof, we must show that $F_W : W \to A$ is a bijection. It is clearly injective. To show that it is surjective, suppose for contradiction that there exists some $b \in A \setminus F_W[W]$. Because W is an ordinal number γ, it satisfies the definition of W (thanks to b) and thus $\gamma = W \in W$, which is a contradiction. □

2.5 Zorn's Lemma

Zorn's lemma is perhaps the most useful application of the axiom of choice outside set theory. It is also known as the maximality principle, a name that goes back to the German mathematician Felix Hausdorff, who proved an earlier and equivalent version of the theorem in 1914 (see [3] for details). The formulation known today as Zorn's lemma was introduced in 1935 by another German mathematician, Max Zorn. However, it had already been independently proven in 1922 by the Polish mathematician Kazimierz Kuratowski, whom you might know for Kuratowski's theorem — a forbidden-graph characterization of planar graphs.

Definition 2.19 (Chain)**.** Let $(a, \leq_R)$ be an ordered set. We call the subset $b \subseteq a$ a *chain* in $a \equiv b$ is totally ordered by $\leq_R$.

Principle 2.20 (Zorn's Lemma)**.** Every (partially) ordered set containing upper bounds for every chain necessarily contains at least one maximal element.

There is also a parametrized version of this statement.

Principle 2.21 (Parametrized Zorn's Lemma)**.** Let A be a (partially) ordered set containing upper bounds for every chain. Then for every $a \in A$, there is a maximal element $b \in A$ such that $a \leq b$.

We can obtain the parametrized version from the unparameterized one by restricting ourselves to the elements above (or equal to) a. The other direction is obvious.

Remark. Zorn's lemma can be made slightly stronger by assuming that only well-ordered chains have upper bounds. The proof remains virtually unchanged.

Theorem 2.22. *The axiom of choice implies Zorn's lemma.*

Proof. Let $(A, <_R)$ be an ordered set containing upper bounds for each chain and for contradiction suppose that there is no maximal element. Note that this implies that every chain in fact has a *strict* upper bound. If a chain C had no strict upper bound, then the non-strict upper bound $b \in C$ would be a maximal element. We denote the set of strict upper bounds of C as $C^>$.

We take $f : \mathcal{P}(A) \to A$, a selector on $\mathcal{P}(A)$, and define a function g from the set of all chains in A as $g(C) := f(C^>)$. So g maps a chain to one of its strict upper bounds. Now pick an arbitrary $a \in A$ and define the mapping $H : \mathrm{On} \to A$ by transfinite recursion as $H(0) = a$ and $H(\alpha + 1) = g(\{H(\alpha)\})$ for successor ordinals, and as $H(\delta) = g(H[\delta])$ for limit ordinals. We start with a and get bigger and bigger elements of A using successor ordinals, each time taking a strict upper bound of a single element chain. If an ordinal δ is limit, we notice that $H[\delta]$ is a chain (all the smaller elements that we picked previously are strict upper bounds of each other and therefore comparable) and $H(\delta)$ is a strict upper bound of this chain.

Note that if we want to be rigorous about the construction by transfinite recursion, we should define g on the entire V. But we can do this in any way, for example by defining $G(x)$ as $\varnothing$ if x is not a chain of A and $g(x)$ otherwise.

Finally, observe that $H : \mathrm{On} \to A$ is an increasing function (each value is a strictly larger upper bound than the previous one) and that it is injective. Thus we obtain an injection from the proper class On into the set A, which is impossible. Indeed, taking the inverse mapping and applying the axiom of replacement would then imply that On itself is a set, which is a contradiction. $\square$

Theorem 2.23. *Zorn's lemma implies the well-ordering principle.*

Proof. Let X be any set. We will find a well-ordering of it by considering all of its possible well-ordered subsets, picking the maximal using Zorn's lemma and showing that it orders the entire X. Consider the set [2]

$$\mathcal{W} \coloneqq \{(A, <_R) \mid <_R \text{ is a well-order on } A \subseteq X\}.$$

We will define a partial order $\preceq_{\mathcal{W}}$ on it like $(A, <_R) \preceq_{\mathcal{W}} (B, <_S) \equiv A \subseteq B$ and $<_R$ is the restriction of $<_S$ to A. We will apply Zorn's lemma to $\mathcal{W}$.

First we need to show that chains have upper bounds. Let $\mathcal{C} \subseteq \mathcal{W}$ be a chain. Define the set

$$M \coloneqq \bigcup \{A \mid (A, <_R) \in \mathcal{C}\} \subseteq X,$$

and for $x, y \in M$ put $x <_M y \equiv \exists (A, <_R) \in \mathcal{C}$ such that $x, y \in A$ and $x <_R y$. Because $\mathcal{C}$ is a chain, this is well defined: if x and y belong to two distinct orderings in $\mathcal{C}$ then one extends the other and hence they agree.

We claim that $(M, <_M)$ is well-ordered. As a consequence of how M is defined, for every nonempty $S \subseteq M$ there exists some well-ordered $(A, <_R) \in \mathcal{C}$ such that $S \subseteq A$. We can use this well-ordering: note that $\min_{<_R}(S) = \min_{<_M}(S)$ because $<_M$ extends $<_R$. Therefore $(M, <_M)$ is well-ordered and thus an upper bound of $\mathcal{C}$ in $\mathcal{W}$.

Because all chains are bounded, by Zorn's lemma, $\mathcal{W}$ has a maximal element $(W, <_W)$. We claim that $W = X$ and so it is the sought-after well-ordering of X. For contradiction, suppose there exists some $x \in X \setminus W$ and extend the ordering $<_W$ to $W' \coloneqq W \cup \{x\}$ by making each $y \in W$ smaller than x. Notice that this slightly "longer" order is a well-ordering of W' and therefore is in $\mathcal{W}$. Moreover, it extends $(W, <_W)$ which hence is not maximal in $(\mathcal{W}, \preceq_{\mathcal{W}})$. We have arrived at a contradiction and can conclude that $W = X$. $\square$

To summarize, in the last few theorems we have shown that the axiom of choice, the well-ordering principle, and Zorn's lemma are all equivalent.

Theorem 2.24. *The following conditions are equivalent in ZF:*

- *the axiom of choice,*
- *the well-ordering principle,*
- *Zorn's lemma.*

---

[2] Why is this a set?

To demonstrate an application of Zorn's lemma, consider the following question. Does every connected graph have a spanning tree? Finding one in a finite graph is easy: simply remove edges of cycles until there are no cycles left. But this process may not terminate for infinite graphs.

Proposition 2.25. *Every connected graph has a spanning tree.*

Proof. The set of all sub-graphs that are trees is partially ordered by inclusion, and the union of a chain is its upper bound. Zorn's lemma says that a maximal tree must exist, which is a spanning tree since the graph is connected. □

Sources

This document is mostly my notes from the class NMAI074 taught at MFF CUNI by doc. Kynčl. The web of the course is HERE. The course mostly follows parts of the second and third chapters of [1], which is in Czech. My notes from the introductory set theory course can be found HERE, also in Czech.

If you found any mistakes or errors, please contact me at smolikj@matfyz.cz.

[1] Petr Balcar Bohuslav a Štěpánek. *Teorie množin.* Vydání 2., opravené a rozšířené. Praha: Academia, 2001. ISBN: 80-200-0470-X.

[2] Karel Hrbáček and Tomáš Jech. *Introduction to set theory.* eng. Third edition, revised and expanded. Pure and applied mathematics. A series of monographs and textbooks ; 220. Boca Raton: Taylor & Francis, 1999. ISBN: 0-8247-7915-0.

[3] Wikipedia. *Hausdorff maximal principle — Wikipedia, The Free Encyclopedia.* `http://en.wikipedia.org/w/index.php?title=Hausdorff%20maximal%20principle&oldid=1300391387`. [Online; accessed 07-October-2025]. 2025.